

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: POTTER, ANDREW A.  
REDMOND, MARK J.  
HUGHES, HUW P.A.

(ii) TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN CHIMERAS

(iii) NUMBER OF SEQUENCES: 11

(iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: ROBERTA L. ROBINS  
(B) STREET: 635 BRYANT STREET  
(C) CITY: PALO ALTO  
(D) STATE: CALIFORNIA  
(E) COUNTRY: UNITED STATES OF AMERICA  
(F) ZIP: 94301

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/960,932  
(B) FILING DATE: 14-OCT-1992  
(C) CLASSIFICATION: 435

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: ROBINS, ROBERTA L.  
(B) REGISTRATION NUMBER: 33,208  
(C) REFERENCE/DOCKET NUMBER: 9000-0016.20

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (415) 617-8999  
(B) TELEFAX: (415) 327-3231

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2794 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..2778

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGCTACTG TTATAGATCT AAGCTTCCCA AAAACTGGGG CAAAAAAAT TATCCTCTAT	60
ATTCCCCAAA ATTACCAATA TGATACTGAA CAGGTAATG GTTTACAGGA TTTAGTCAAA	120
GCGGCCGAAG AGTTGGGAT TGAGGTACAA AGAGAAGAAC GCAATAATAT TGCAACAGCT	180

DRAFT

CAAACCAGTT TAGGCACGAT TCAAACCGCT ATTGGCTTAA CTGAGCGTGG CATTGTGTTA	240
TCCGCTCCAC AAATTGATAA ATTGCTACAG AAAACTAAAG CAGGCCAACG ATTAGGTTCT	300
GCCGAAAGCA TTGTACAAAA TGCAAATAAA GCCAAAAGT GATTATCTGG CATTCAATCT	360
ATTTTAGGCT CAGTATTGGC TGGAATGGAT TTAGATGAGG CCTTACAGAA TAACAGCAAC	420
CAACATGCTC TTGCTAAAGC TGGCTGGAG CTAACAAATT CATTAAATTGA AAATATTGCT	480
AATTCAAGTAA AAACACTTGA CGAATTGGT GAGCAAATTA GTCAATTGG TTCAAAACTA	540
CAAAATATCA AAGGCTTAGG GACTTTAGGA GACAAACTCA AAAATATCGG TGGACTTGAT	600
AAAGCTGGCC TTGGTTAGA TGTTATCTCA GGCTATTAT CGGGCGCAAC AGCTGCACTT	660
GTACTTGCAG ATAAAAATGC TTCAACAGCT AAAAAAGTGG GTGCGGGTT TGAATTGGCA	720
AACCAAGTTG TTGGTAATAT TACCAAGCC GTTCTTCTT ACATTTCAGC CCAACGTGTT	780
GCAGCAGGTT TATCTTCAAC TGGGCCTGTG GCTGCTTAA TTGCTTCTAC TGTTCTCTT	840
GCGATTAGCC CATTAGCATT TGCCGGTATT GCGATAAT TTAATCATGC AAAGAGTTA	900
GAGAGTTATG CCGAACGCTT TAAAAAATTA GGCTATGACG GAGATAATT ATTAGCAGAA	960
TATCAGCGGG GAACAGGGAC TATTGATGCA TCGGTTACTG CAATTAAATAC CGCATTGGCC	1020
GCTATTGCTG GTGGTGTGTC TGCTGCTGCA GCCGGCTCGG TTATTGCTTC ACCGATTGCC	1080
TTATTAGTAT CTGGGATTAC CGGTGTAATT TCTACGATTC TGCAATATTC TAAACAAAGCA	1140
ATGTTTGAGC ACGTTGCAA AAAAAATTCA AACAAATATG TAGAATGGGA AAAAAATAAT	1200
CACGGTAAGA ACTACTTTGA AAATGGTAC GATGCCCGTT ATCTTGCAGA TTTACAAGAT	1260
AATATGAAAT TCTTACTGAA CTTAAACAAA GAGTTACAGG CAGAACGTGT CATCGCTATT	1320
ACTCAGCAGC AATGGGATAA CAACATTGGT GATTTAGCTG GTATTAGCCG TTTAGGTGAA	1380
AAAGTCCTTA GTGGTAAAGC CTATGTGGAT GCGTTGAAG AAGGCAAAACA CATTAAAGCC	1440
GATAAAATTAG TACAGTTGGA TTCGGCAAAC GGTATTATTG ATGTGAGTAA TTGGGTAAA	1500
GCGAAAACTC AGCATATCTT ATTCAAGAACG CCATTATTGA CGCCGGAAC AGAGCATCGT	1560
GAACCGTAC AAACAGGTAA ATATGAATAT ATTACCAAGC TCAATATTAA CCGTGTAGAT	1620
AGCTGGAAAA TTACAGATGG TGCAGCAAGT TCTACCTTG ATTTAACTAA CGTTGTTCA	1680
CGTATTGGTA TTGAATTAGA CAATGCTGGA AATGTAACCA AAACCAAAGA AACAAAAATT	1740
ATTGCCAAC TTGGTGAAGG TGATGACAAC GTATTTGTTG GTTCTGGTAC GACGGAAATT	1800
GATGGCGGTG AAGGTTACGA CCGAGTTCAC TATAGCCGTG GAAACTATGG TGCTTTAACT	1860
ATTGATGCAA CCAAAGAGAC CGAGCAAGGT AGTTATACCG TAAATCGTT CGTAGAAACC	1920
GGTAAAGCAC TACACGAAGT GACTTCAACC CATAACCGCAT TAGTGGCAA CCGTGAAGAA	1980
AAAAATAGAAT ATCGTCATAG CAATAACCGAG CACCATGCCG GTTATTACAC CAAAGATAAC	2040

TTGAAAGCTG TTGAGAAT TATCGTACA TCACATAACG ATATCTTAA AGGTAGTAAG 2100  
TTCAATGATG CCTTTAACGG TGGTGATGGT GTCGATACTA TTGACGGTAA CGACGGCAAT 2160  
GACCGCTTAT TTGGTGGTAA AGGCATGAT ATTCTCGATG GTGGAAATGG TGATGATT 2220  
ATCGATGGCG GTAAAGGCAA CGACCTATTA CACGGTGGCA AGGGCGATGA TATTTCGTT 2280  
CACCGTAAAG GCGATGGTAA TGATATTATT ACCGATTCTG ACGGCAATGA TAAATTATCA 2340  
TTCTCTGATT CGAACTTAAA AGATTAACA TTTGAAAAAG TTAAACATAA TCTTGTCA 2400  
ACGAATAGCA AAAAGAGAA AGTGACCATT CAAAAGTGGT TCCGAGAGGC TGATTTGCT 2460  
AAAGAAGTGC CTAATTATAA AGCAACTAAA GATGAGAAAA TCGAAGAAT CATCGTCAA 2520  
AATGGCGAGC GGATCACCTC AAAGCAAGTT GATGATCTTA TCGCAAAAGG TAACGGCAA 2580  
ATTACCCAAG ATGAGCTATC AAAAGTTGTT GATAACTATG AATTGCTCAA ACATAGCAA 2640  
AATGTGACAA ACAGCTTAGA TAAGTTAATC TCATCTGTAA GTGCATTTAC CTCGTCTAAT 2700  
GATTCGAGAA ATGTATTAGT GGCTCCACT TCAATGTTGG ATCAAAGTTT ATCTTCTCTT 2760  
CAATTGCTA GGGGATCCTA GCTAGCTAGC CATG 2794

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 60 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GATCCAGCTC TTCTGCCGGC TGCAAAACT TCTTCTGGAA AACCTTCACC AGCTGCTAGG 60

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 60 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GATCCCTAGC AGCTGGTGAA GGTTTCCAG AAGAAGTTTT TGCAGCCGGC AGAAGAGCTG 60

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

2025 RELEASE UNDER E.O. 14176

(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:	
GATCTCAGCA TTGGAGCTAC GGCGCTGCGCC CTGGCTAAG	39
(2) INFORMATION FOR SEQ ID NO:5:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 39 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
GATCCTTAGC CAGGGCGCAG GCCGTAGCTC CAATGCTGA	39
(2) INFORMATION FOR SEQ ID NO:6:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 83 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
GATCTGCCTA CATTGTGCCT GTGAGCATTG TGAGCCGCAA CATTGTGTAC ACCCGCCGCG	60
AACCTAACCA AGACATTGTG TAG	83
(2) INFORMATION FOR SEQ ID NO:7:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 83 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:	
GATCCTACAC AATGTCTTGG TTAAGTTGCG CGCGGGTGTA CACAATGTTG CGGCTCACAA	60
TCGTCACAGG CACAATGTTG CAA	83
(2) INFORMATION FOR SEQ ID NO:8:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 2838 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	

## (ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 1..2829

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGGCTACTG TTATAGATCT AAGCTTCCCA	AAAAACTGGGG CAAAAAAAT TATCCTCTAT	60
ATTCCCCAAA ATTACCAATA TGATACTGAA	CAAGGTAAATG GTTTACAGGA TTTACTCAA	120
CGGGCCGAAG AGTTGGGAT TGAGGTACAA	AGAGAAGAAC GCAATAATAT TGCAACAGCT	180
CAAACCAGTT TAGGCACCGAT	TCAAAACCGCT ATTGGCTTAA CTGAGCGTGG CATTGTGTTA	240
TCCGCTCCAC AAATTGATAA ATTGCTACAG	AAACTAAAG CAGGCCAGC ATTAGGTCT	300
GCCGAAAGCA TTGTACAAA TGCATAAAA	GCCAAAAC TG TATTATCTGG CATTCAATCT	360
ATTTTAGGCT CAGTATTGGC	TGGAATGGAT TTAGATGAGG CCTTACAGAA TAACAGCAAC	420
CAACATGCTC TTGCTAAAGC	TGGCTTGGAG CTAACAAATT CATTAAATTGA AAATATTGCT	480
AATTCACTAA AAACACTTGA CGAATTGGT	GAGCAAATT GA TC AATTGG TTCAAAACTA	540
CAAATATCA AAGGCTTAGG	GA CTTAGGA GACAAACTCA AAAATATCGG TGGACTTGAT	600
AAAGCTGGCC TTGGTTAGA	TGTTATCTCA GGGCTATTAT CGGGCGCAG AGCTCCACTT	660
GTACTTGCAG ATAAAAATGC	TTCAACAGCT AAAAAAGTGG GTGCGGGTTT TGAATTGGCA	720
AACCAAGTTG TTGGTAATAT	TACCAAAGCC GPTTCTTCTT ACATTTAGC CCAACGTGTT	780
GCAGCAGGTT TATCTTCAC	TGGGCCTGTG GCTGCTTAA TTGCTTCTAC TGTCTCTTT	840
GCGATTAGCC CATTAGCATT	TGCGGTATT GCGATTAAT TTAATCATGC AAAAAGTTA	900
GAGACTTATG CCGAACGCTT	TAAATTA GGCTATGACG GAGATAATT ATTAGCAGAA	960
TATCAGCGGG	GAACAGGGAC TATTGATGCA TCGGTTACTG CATTAAATAC CGCATTGCC	1020
GCTATTGCTG	GTGGTGTGTC TGCTGCTGCA GCCGGCTCGG TTATTGCTTC ACCGATTGCC	1080
TTATTAGTAT	CTGGGATTAC CGGTGTATT TCTACGATTC TGCAATATTC TAAACAAAGCA	1140
ATGTTGAGC	ACGTTGAAA TAAATTCAT AACAAAATG TAGAATGGGA AAAAATAAT	1200
CACGGTAAGA	ACTACTTTGA AAATGCTTAC GATGCCCGTT ATCTTGCAGA TTTACAAGAT	1260
AATATGAAAT	TCTTACTGAA CTTAAACAAA GAGTTACAGG CAGAACGTGT CATCGCTATT	1320
ACTCAGCAGC	AATGGGATAA CAACATTGGT GATTAGCTG GTATTAGCCG TTTAGGTGAA	1380
AAAGTCCTTA	GTGGTAAAGC CTATGTGGAT GCGTTTGAAG AAGGCAAACA CATTAAAGCC	1440
GATAAAATTAG	TACAGTTGGA TTCGGCAAAAC GGTATTATTG ATGTGAGTAA TTCGGGTAAA	1500
GCGAAAACTC	AGCATATCTT ATTCAAGAACG CCATTATTGA CGCCGGAAAC AGAGCATCGT	1560
GAACGCGTAC	AAACAGGTAA ATATGAATAT ATTACCAAGC TCAATATTAA CGGTGTAGAT	1620
AGCTGGAAAAA	TTACAGATGG TGCAGCAAGT TCTACCTTG ATTTAACTAA CGTTGTTCA	1680

CGTATTGGTA TTGAATTAGA CAATGCTGGA AATGTAACTA AAACCAAAGA AACAAAAATT 1740  
 ATTGCCAAC TTGGTGAAGG TGATGACAAC GTATTTGTTG GTTCTGGTAC GACGGAAATT 1800  
 GATGGCGGTG AAGGTTACGA CCGAGTTCAC TATAGCCGTG GAAACTATGG TGCTTTAACT 1860  
 ATTGATGCAA CCAAAGAGAC CGAGCAAGGT AGTTATAACCG TAAATCGTTT CGTAGAAACC 1920  
 GGTAAAGCAC TACACGAAGT GACTTCACC CATACCGCAT TAGTGGGCAA CCGTGAAGAA 1980  
 AAAATAGAAT ATCGTCATAG CAATAACCAG CACCATGCCG GTTATTACAC CAAAGATAACC 2040  
 TTGAAAGCTG TTGAAGAAAT TATCGGTACA TCACATAACG ATATCTTTAA AGGTAGTAAG 2100  
 TTCAATGATG CCTTTAACCG TGGTGATGGT GTCGATACTA TTGACGGTAA CGACGGCAAT 2160  
 GACCGCTTAT TTGGTGGTAA AGGCATGAT ATTCTCGATG GTGGAAATGG TGATGATTT 2220  
 ATCGATGGCG GTAAAGGCAA CGACCTATTA CACGGTGGCA AGGGCGATGA TATTTCGTT 2280  
 CACCGTAAAG GCGATGGTAA TGATATTATT ACCGATTCTG ACGGCAATGA TAAATTATCA 2340  
 TTCTCTGATT CGAACTTAAA AGATTTAACCA TTTGAAAG TTAAACATAA TCTTGTCACTC 2400  
 ACGAATAGCA AAAAAGAGAA AGTGACCAATT CAAACTGGT TCCGAGAGGC TGATTTGCT 2460  
 AAAGAAGTGC CTAAATTATAA AGCAACTAAA GATGAGAAA TCGAAGAAAT CATCGGTCAA 2520  
 AATGGCGAGC GGATCACCTC AAAGCAAGTT GATGATCTTA TCGAAAAGG TAACGGCAA 2580  
 ATTACCCAAG ATGAGCTATC AAAAGTTGTT GATAACTATG AATTGCTCAA ACATAGCAA 2640  
 AATGTGACAA ACAGCTTAGA TAAGTTAACC TCATCTGAA GTGCATTTAC CTCGTCTAA 2700  
 GATTGAGAA ATGTATTAGT GGCTCCAACT TCAATGTTGG ATCAAGTTT ATCTTCTCTT 2760  
 CAATTTGCTA GGGGATCCAG CTCTTCTGCC GGCTGCAAAA ACTTCTTCTG GAAAACCTTC 2820  
 ACCAGCTGCT AGGGATCC 2838

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2817 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2808

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATGGCTACTG TTATAGATCT AAGCTTCCCA AAAACTGGGG CAAAAAAAT TATCCTCTAT 60  
 ATTCCCCAAA ATTACCAATA TGATACTGAA CAGGTAAATG GTTTACAGGA TTTAGTCAAA 120  
 CGGGCCGAAG AGTTGGGAT TGAGGTACAA AGAGAAGAAC GCAATAATAT TGCAACAGCT 180

CAAAACCAGTT TAGGCACCGAT TCAAAACCGCT ATTGGCTTAA CTGAGCGTGG CATTGTGTTA  
TCCGCTCCAC AAATTGATAA ATTGCTACAG AAAACTAAAG CAGGCCAAGC ATTAGGTTCT 300  
GCCGAAAGCA TTGTACAAAA TGCAAATAAA GCCAAAATG TATTATCTGG CATTCAATCT 350  
ATTTAGGCT CAGTATTGGC TGGAATGGAT TTAGATGAGG CCTTACAGAA TAACAGCAAC 400  
CAACATGCTC TTGCTAAAGC TGGCTGGAG CTAACAAATT CATTAAATTGA AAATATTGCT 450  
AATTCAAGTAA AAACACTTGA CGAATTGGT GAGCAAATT GTCAATTGG TTCAAAAC 500  
CAAATATCA AAGGCTTAGG GACTTTAGGA GACAACCTCA AAAATATCGG TGGACTTGAT 550  
AAAGCTGGCC TTGGTTAGA TGTTATCTCA GGGCTATTAT CGGGCGCAAC AGCTGCAC 600  
GTACTTGCAG ATAAAAATGC TTCAACAGCT AAAAAAGTGG GTGCGGGTTT TGAATTGGCA 650  
AACCAAGTTG TTGGTAATAT TACCAAAGCC GTTTCTCTT ACATTTAGC CAAACGTGTT 700  
GCAGCAGGTT TATCTTCAAC TGGGCCTGTG GCTGCTTTAA TTGCTTCTAC TGGTTCTCTT 750  
GCGATTAGCC CATTAGCATT TGCGGTATT GCCGATAAAAT TTAATCATGC AAAAAAGTTA 800  
GAGAGTTATG CCGAACGCTT TAAAAAATTA GGCTATGACG GAGATAATT ATTAGCAGAA 850  
TATCAGCGGG GAACAGGGAC TATTGATGCA TCGGTTACTG CAATTAATAC CGCATTGGCC 900  
GCTATTGCTG TGCGGTGTGTC TGCTGCTGCA GCCGGCTCGG TTATTGCTTC ACCGATTGCC 950  
TTATTAGTAT CTGGGATTAC CGGTGTAATT TCTACGATT TCAGATATTCAAAACAAGCA 1000  
ATGTTTGGC ACGTTGCAAA TAAAATTCA AAAAAATTG TAGAATGGGA AAAAAATAAT 1050  
CACGGTAAGA ACTACTTGA AATGGTTAC GATGCCCGTT ATCTTGCAGA TTTACAAGAT 1100  
AATATGAAAT TCTTACTGAA CTTAAACAAA GAGTTACAGG CAGAACGTGT CATCGCTATT 1150  
ACTCAGCAGC AATGGGATAA CAACATTGGT GATTTAGCTG GTATTAGCCG TTTAGGTGAA 1200  
AAAGTCCTTA GTGGTAAAGC CTATGTGGAT GCGTTGAAAG AAGGCAAACA CATTAAAGCC 1250  
GATAAATTAG TACAGTTGCA TTCCGGCAAAAC GGTATTATTG ATGTGAGTAA TCGGGTAAA 1300  
GCGAAAACTC AGCATATCTT ATTCAAGAACG CCATTATTGA CGCCGGAAC AGAGCATCGT 1350  
GAGCGCGTAC AAACAGGTAA ATATGAATAT ATTACCAAGC TCAATATTAA CCGTGTAGAT 1400  
AGCTGGAAAA TTACAGATGG TGCAGCAAGT TCTACCTTG ATTTAACTAA CGTTGTTCA 1450  
CGTATTGGTA TTGAATTAGA CAATGCTGGA AATGTAACCA AAACCAAAGA AACAAAAATT 1500  
ATTGCCAAAC TTGGTGAAGG TGATGACAAC GTATTTGTTG GTTCTGGTAC GACGGAAATT 1550  
GATGGCCGTG AAGGTTACGA CCGAGTTACAC TATAGCCGTG GAAACTATGG TGCTTTAACT 1600  
ATTGATGCAA CCAAAGAGAC CGAGCAAGGT AGTTATACCG TAAATCGTTT CGTAGAAACC 1650  
GGTAAAGCAC TACACGAAGT GACTTCAACC CATACCGCAT TACTGGCAA CCGTGAAGAA 1700  
AAAATAGAAT ATCGTCATAG CAATAACCAG CACCATGCCG GTTATTACAC CAAAGATAAC 1750  
2040

TTGAAAGCTG TTGAAGAAAT TATCGGTACA TCACATAACG ATATCTTAA AGGTAGTAAG	2100
TTCAATGATG CCTTTAACGG TGGTGATGGT GTCGATACTA TTGACGGTAA CGACGGCAAT	2160
GACCGCTTAT TTGGTGGTAA AGGCGATGAT ATTCTCGATG GTGGAAATGG TGATGATTT	2220
ATCGATGGCG GTAAAGGCAA CGACCTATTA CACGGTGGCA AGGGCGATGA TATTTCGTT	2280
CACCGTAAAG GCGATGGTAA TGATATTATT ACCGATTCTG ACGGCAATGA TAAATTATCA	2340
TTCTCTGATT CGAACTTAAA AGATTTAAC A TTTGAAAG TTAAACATAA TCTTGTCA	2400
ACGAATAGCA AAAAAGAGAA AGTGACCATT CAAAACTGGT TCCGAGAGGC TGATTTGCT	2460
AAAGAAGTGC CTAATTATAA AGCAACTAAA GATGAGAAAA TCGAAGAAAT CATCGGTCAA	2520
AATGGCGAGC GGATCACCTC AAAGCAAGTT GATGATCTTA TCGCAAAAGG TAACGGCAA	2580
ATTACCCAAG ATGAGCTATC AAAAGTTGTT GATAACTATG AATTGCTCAA ACATAGCAA	2640
AATGTGACAA ACAGCTTAGA TAAGTTAATC TCATCTGTAA GTGCATTTAC CTCGTCTAAT	2700
GATTGAGAA ATGTATTAGT GGCTCCA ACT TCAATGTTGG ATCAAAGTT ATCTTCTCTT	2760
CAATTTGCTA GGGGATCTCA GCATTGGAGC TACGGCCTGC GCCCTGGCTA AGGATCC	2817

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2861 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..2853

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGGCTACTG TTATAGATCT AAGCTTCCC AAAACTGGGG CAAAAAAAT TATCCTCTAT	60
ATTCCCCAAA ATTACCAATA TGATACTGAA CAAGGTAAATG GTTACAGGA TTTAGTCAAA	120
GCAGGCCAGAG AGTTGGGGAT TGAGGTACAA AGAGAAGAAC GCAATAATAT TGCAACAGCT	180
CAAACCAGTT TAGGCACGAT TCAAACCGCT ATTGGCTTAA CTGAGCGTGG CATTGTGTTA	240
TCCGCTCCAC AAATTGATAA ATTGCTACAG AAAACTAAAG CAGGCCAACG ATTAGTTCT	300
GCCGAAAGCA TTGTACAAAA TGCAAATAAA GCCAAAATG TATTATCTGG CATTCAATCT	360
ATTTTAGGCT CAGTATTGGC TGGAAATGGAT TTAGATGAGG CCTTACAGAA TAACAGCAAC	420
CAACATGCTC TTGCTAAAGC TGGCTTGGAG CTAACAAATT CATTAAATTGA AAATATTGCT	480
AATTCAAGTAA AAACACTTGA CGAATTTGGT GAGCAAATTA GTCAATTGG TTCAAAACTA	540
CAAAATATCA AAGGCTTAGG GACTTTAGGA GACAAACTCA AAAATATCGG TGGACTTGAT	600

AAAGCTGGCC TTGGTTAGA TGTTATCTCA GGGCTATTAT CGGGCGAAC AGCTGCAC	650
TTGCAAG ATAAAAATGC TTCAACAGCT AAAAAAGTGG GTGCCGGTTT TGAATTGGCA	720
AACCAAGTTG TTGGTAATAT TACCAAAGCC GTTTCTTCTT ACATTTAGC CCAACGTGTT	780
GCAGCAGGTT TATCTTCAAC TGGGCCTGTG GCTGCTTAA TTGCTTCTAC TGTTCTCTT	840
GCGATTAGCC CATTAGCATT TGCCGGTATT GCCGATAAAAT TTAATCATGC AAAAAAGTTA	900
GAGAGTTATG CCGAACGCTT TAAAAAATTA GGCTATGACG GAGATAATTT ATTAGCAGAA	960
TATCAGCGGG GAACAGGGAC TATTGATGCA TCGGTTACTG CAATTAATAC CGCATTGCC	1020
GCTATTGCTG GTGGTGTGTC TGCTGCTGCA GCCGGCTCGG TTATTGCTTC ACCGATTGCC	1080
TTATTAGTAT CTGGGATTAC CGGTGTAATT TCTACGATT TGCAATATTC TAAACAAGCA	1140
ATGTTGAGC ACGTTGAAA TAAAATTCA TACAAAATTG TAGAATGGGA AAAAAATAAT	1200
CACGGTAAGA ACTACTTGAA AAATGGTTAC GATGCCCGTT ATCTTGCAGA TTTACAAGAT	1260
AATATCAAAT TCTTACTGAA CTTAACAAA GAGTTACAGG CAGAACGTGT CATCGCTATT	1320
ACTCAGCAGC AATGGGATAA CAACATTGGT GATTTAGCTG GTATTAGCCG TTTAGGTGAA	1380
AAAGTCCTTA GTGGTAAAGC CTATGTGGAT GCGTTGAG AAGGCAAACA CATTAAAGCC	1440
GATAAATTAG TACAGTTGGA TTCGGCAAAAC GGTATTATTG ATGTGAGTAA TTGGGTAAA	1500
GCGAAAACTC AGCATATCTT ATTCAAGAACG CCATTATTGA CGCCGGAAC AGAGCATCGT	1560
GAACCGCTAC AAACAGGTAA ATATGAATAT ATTACCAAGC TCAATATTAA CCGTGTAGAT	1620
AGCTGGAAAA TTACAGATGG TGCAGCAAGT TCTACCTTG ATTTAACTAA CGTTGTTCA	1680
CGTATTGGTA TTGAATTAGA CAATGCTGGA AATGTAACCA AAACCAAAGA AACAAAAATT	1740
ATTGCCAAC TTGGTGAAGG TGATGACAAC GTATTGTTG GTTCTGGTAC GACGGAAATT	1800
GATGGCGGTG AAGGTTACGA CCGAGTTAC TATAGCCGTG GAAACTATGG TGCTTTAAGT	1860
ATTGATGCCAA CCAAAGAGAC CGAGCAAGGT AGTTATACCG TAAATCGTTT CGTAGAAACC	1920
GGTAAAGCAC TACACGAAGT GACTTCAACC CATAACCGCAT TAGTGGCAA CCGTGAAGAA	1980
AAAATAGAAT ATCGTCATAG CAATAACCAG CACCATGCCG GTTATTACAC CAAGATACC	2040
TTGAAAGCTG TTGAAGAAAT TATCGGTACA TCACATAACG ATATCTTAA AGGTAGTAAG	2100
TTCAATGATG CCTTTAACGG TGGTGTGGT GTCGATACTA TTGACGGTAA CGACGGCAAT	2160
GACCGCTTAT TTGGTGGTAA AGGCAGATGAT ATTCTCGATG GTGGAAATGG TGATGATTT	2220
ATCGATGGCG GTAAAGGCAA CGACCTATTA CACGGTGGCA AGGGCGATGA TATTTCGTT	2280
CACCGTAAAG GCGATGGTAA TGATATTATT ACCGATTCTG ACGGCAATGA TAAATTATCA	2340
TTCTCTGATT CGAACTTAAA AGATTTAACCA TTTGAAAAAG TTAAACATAA TCTTGTCA	2400
ACGAATAGCA AAAAGAGAA AGTGACCATT CAAACTGGT TCCGAGAGGC TGATTTGCT	2460

AAAGAAGTGC CTAATTATAA AGCAACTAAA GATGAGAAAA TCGAAGAAAT CATCGGTCAA 2520  
AATGGCGAGC GGATCACCTC AAAGCAAGTT GATGATCTTA TCGCAAAAGG TAACGGCAAA 2580  
ATTACCCAAG ATGAGCTATC AAAAGTTGTT GATAACTATG AATTGCTCAA ACATAGCAAA 2640  
AATGTGACAA ACAGCTTAGA TAAGTTAAC TCATCTGTAA GTGCATTTAC CTCGTCTAAT 2700  
GATTGAGAA ATGTATTAGT GGCTCCAATC TCAATGTTGG ATCAAAGTTT ATCTTCTCTT 2760  
CAATTTGCTA GGGGATCTTG CAACATTGTG CCTGTGAGCA TTGTGAGCCG CAACATTGTG 2820  
TACACCCCGCG CGCAACCTAA CCAAGACATT GTGTAGGATC C 2861

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 3
- (D) OTHER INFORMATION: /note= "The amino acid at this location can be either Lys, Asp, Val or Asn."

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /note= "The amino acid at this location can be either Lys, Asp, Val or Asn."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Gly Gly Xaa Gly Xaa Asp  
1 5